Approaches to Data Analysis and Presentation

(DRAFT)

Prepared for the project "Remote Sensing Applications for Environmental Analysis in Transportation Planning: Application to the Washington State I-405 Corridor"

Demin Xiong, Russell Lee and J. Bo Saulsbury
Oak Ridge National Laboratory

Elizabeth Lanzer and Albert Perez
Washington State Department of Transportation

August 2002

1. Introduction

Data analysis and presentation is an important task of the project to develop linkages among land use and land cover (LULC) maps, transportation infrastructure and the environment. The task is to fuse data that are acquired or created in previous tasks to analyze and illustrate relationships between transportation infrastructure and the social and natural environments along the I-405 corridor. The task will create maps of overlays of LULC, transportation networks and population distributions and will generate statistics that will serve as indicators of environmental impact for the proposed road infrastructure.

As illustrated by Figure 1, data analysis and presentation represent one of the two major data processing tasks of the project. The first data processing task involves image data processing for LULC classification. After LULC classification, data analysis and presentation procedure follows, which is to provide more relevant information for environmental assessment. Consequently, LULC maps generated from previous image processing task become the primary input for the data analysis and presentation process. To facilitate data analysis and presentation, several additional data sets will be utilized. These include transportation networks derived from Census TIGER files, hydro-networks that already exist in WSDOT, Census population data, social economic data and habitat data. Image data and Digital Elevation Model (DEM) will be part of the input in order to render realistic visualizations of the study area.

The task of data analysis and presentation makes use of several simple, but powerful tools to analyze and present data. These tools include spatial data overlay, buffering, change comparison and generation of 2-D and 3-D perspective displays. Spatial data overlay involves both raster data overlay and vector data overlay, representing a dominant approach to data analysis. The LULC map, once derived, will be in a raster format. To example the relationship between LULC classes and road network infrastructure, a compatible raster data layer of road networks will be constructed and then the raster LULC map then is overlaid on the road network map to show the correlation of these two feature layers. In contrast, relationships between road networks and population distributions or between road networks and hydro-networks will be resolved in the vector domain.

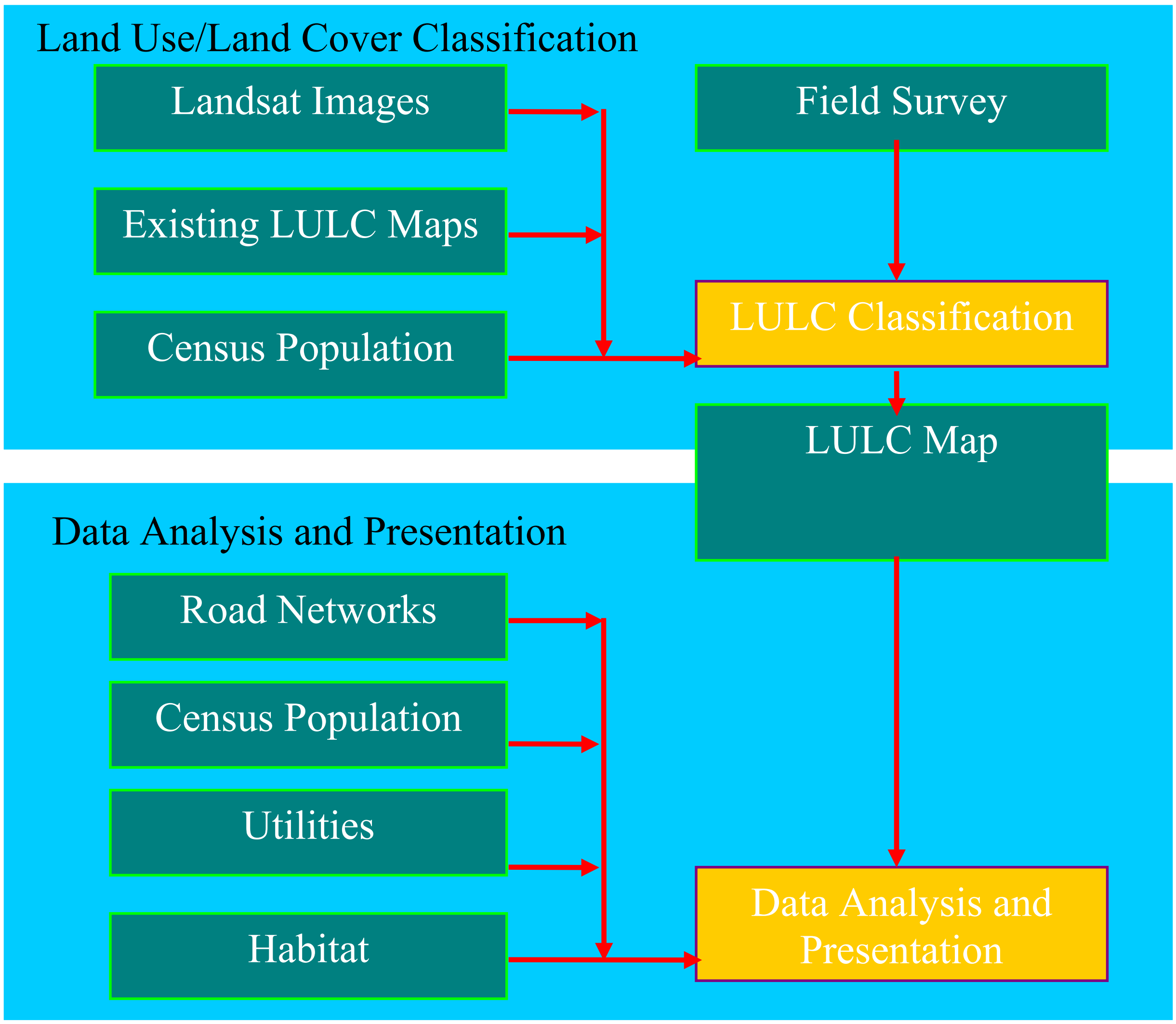

Figure 1. Land Use/Land Cover Classification and Data Analysis and Presentation.

Buffering is another major approach to studying the potential impact of transportation systems to social and natural environments. That is, selected road segments can be used to generate buffer zones and these buffer zones then can be overlaid with feature layers, such as land use and land cover classes and population maps, to obtain indicators of potential impacts that the selected road segments may exert on the environments. The results of buffering can be either in graphical representations (e.g., maps) or in tabular forms (e.g., statistics).

To anticipate and analyze changes of environmental conditions along the road infrastructure, historical land use data, data of population distributions in different time periods, social economic data in multi-years have been collected. A comparison of these data in different time slots will be conducted to identify trends of the changes of the social and natural environments. Increase or decrease regarding particular land use categories along the corridor will be identified as well as changes of population distributions and social economic characteristics. The proposed analysis will stop short from predicting or forecasting changes, but will provide specific indicators of these changes during selected time intervals.

The results of the data analysis will be presented in several different formats. First, the overall analysis results will be included in a database, this database will be accessible with the use of GIS or image processing software such as ESRI's ArcView or ERDAS IMAGINE. Second, softcopies and hardcopies of image displays or GIS maps will be prepared for important data items, e.g., land use maps or color images along the corridor. Third, statistical data will be derived to provide social and environmental characteristics along the corridor (e.g., land use types and areas or population distributions or household numbers) and the changes of these characteristics.

2. Data Analysis

The primary goal of the project is to identify and classify land use and land cover (LULC) – basic information needed in transportation-related environmental analysis and planning. The current task is particularly designed to utilize the LULC data derived from previous tasks together with several other data sets to develop linkages among LULC, transportation infrastructure and some other social and natural environmental factors. The purpose is to generate specific transportation-related environmental impact information to facilitate environmental impact assessment. As LULC represents a crucial element among many other environmental factors and the LULC data represents the single most important product of the project, LULC data analysis becomes the major focus of the data analysis task.

The planned method used for LULC characterization along the I-405 corridor would be a combination of buffering and overlay, which are classical spatial analysis tools in GIS. For this analysis, the input data will include LULC maps and road networks. Type types of LULC maps will be utilized. The first type of map is the map derived from the imaging analysis, representing the current LULC characteristics along the corridor, which will be in a raster data format. The second type of map is the map that comes from USGS, representing LULC characteristics about 20 years ago for the same area, which is in a vector data format. To characterize LULC along the proposed road infrastructure, the referenced road segments will be first identified from the road network map, and then buffers will be generated along these selected road segments. Because of the differences in raster and vector presentation, the buffers will be generated in both raster and vector formats.

In the raster format, cells in different buffer zones or within different distance from the selected road segments will be labeled with corresponding coding values, then the raster LULC map is overlaid on top of these buffers. Then the LULC map is re-coded with a buffer label attached to each of the LULC categories. To do so, LULC categories along the information of proximity to the road infrastructure are obtained. In the vector format, buffer zones that have different distances from the selected road segments are represented as polygons. The vector LULC will also be represented as polygons. The polygon overlays of the vector road buffers and the vector

LULC map will generate a new LULC layer on which LULC categories are tagged with buffer labels, indicating proximity to the selected road segments. Figure 2 shows road buffers overlaid on the USGS LULC map.

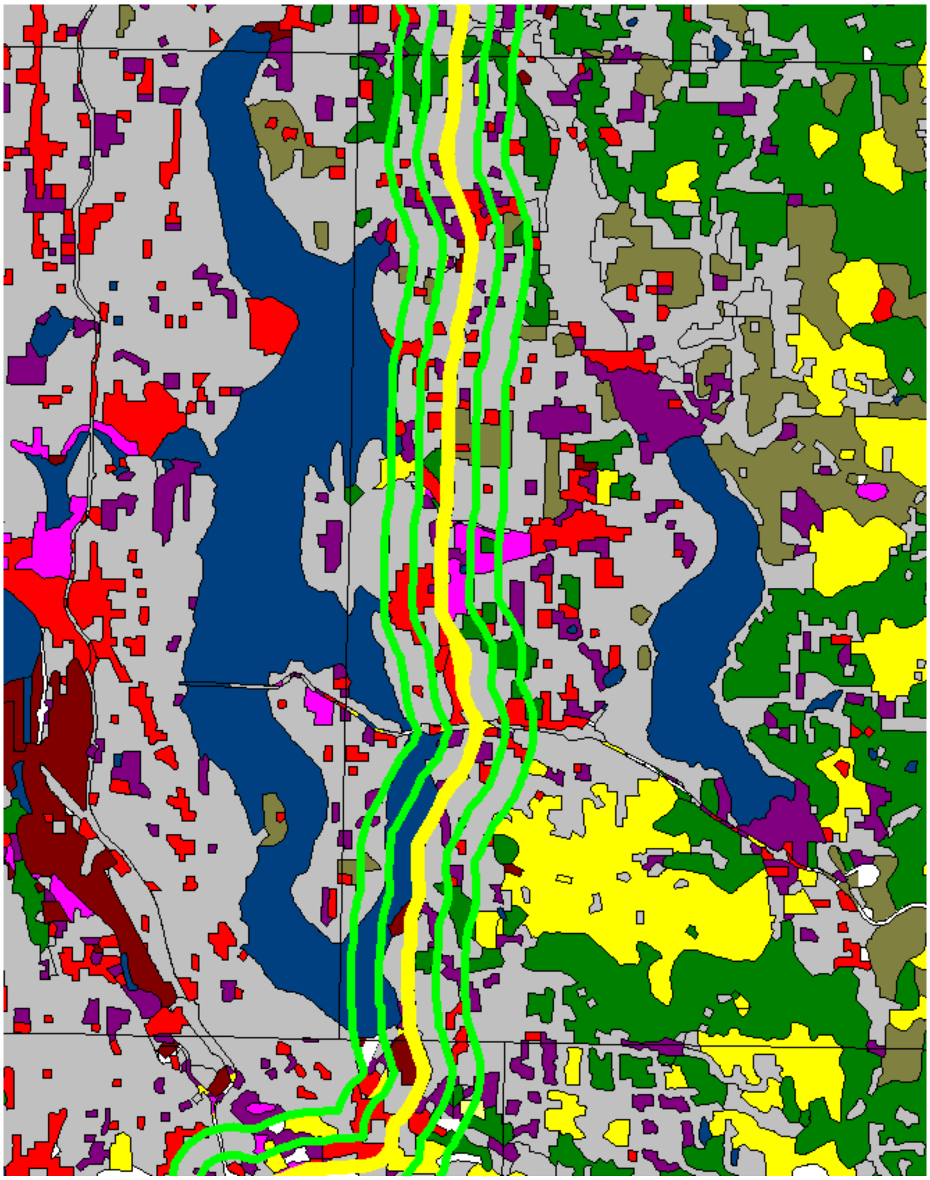

Figure 2. Buffer boundaries, shown as green, with distances of 0.5 and 1 miles from I-405, shown as yellow, overlaid on the USGS LULC map.

After buffering and overlay, statistics of LULC categories for given distances from the selected road segments can be established. Particularly, land areas of any given LULC categories with a given distance from the roads will be computed and recorded, which serve as an indicator of the LULC that may be under influence of the road infrastructure. The statistical difference between buffer and overlay analysis of each LULC map will indicate areas that have changed. Alternatively, the vector LULC map can be converted into a raster map, and then the two LULC maps will be overlaid to run a cell-by-cell comparison. The result of this comparison will provide land use and land cover change or conversion on each cell. When this result is overlaid with the road buffer layer, changes along the selected road segments can be measured accurately.

The general approach used for LULC analysis can be extended for analysis of population distributions, habitats, hydro-networks or water bodies and social economics. Population distributions of Year 2000 and Year 1990 that come from U.S. Census bureau will be overlaid with the road buffers to establish a population and household account along the corridor. Given the differences between year 2000 and year 1990, changes and trends of changes of population distributions will then be identified, displayed and tabulated. Noticeably, population distributions are generally represented at the block level with blocks represented as polygons. Given the spatial detail of the data, it is also feasible to use block centroids to represent population distributions at the block level so that population and households can be counted by simply identifying whether a centroid is located inside a given buffer, which will significantly simplify the counting process. Habitats can be represented with points, lines or polygons. By overlaying habitat layers with road buffers, the area, length or number of the habitats in the proximity to the roads can be computed and recorded.

Hydro-networks and/or water bodies that are represented as lines or polygons can be directly overlaid with the selected road segments or buffers of the selected road segments. In the first case, the length of road segments that cut through water bodies will be measured as well as the intersections between roads and hydro-networks will be identified and counted. In the second case, the length of shore lines and the area of the water bodies and the length of the hydro-networks that are located in the proximity to the selected road segments will be measured and recorded. Both are then can be used as indicators of the impact of the proposed infrastructure on hydro-systems. The social economic data come at the level of the zip code areas in different periods of times. Similar processing method to population distributions will be used to analyze social economic distributions and changes along the corridor.

3. Data Presentation

A major product of image processing and data analysis of the project will be a spatial database which will contain remotely sensed imagery (e.g., color orthophotos, black and white imagery, and Landsat ETM+ data) and Geographical Information system (GIS) data layers (e.g., road networks, demographic data, and digital elevation model (DEM)). The database will also contain LULC maps along with analysis results such as road segment buffering and overlay for LULC characteristics along the I-405 corridor. The database will be constructed in a way that will allow easy access through the use of the GIS or image processing software such as ArcView or IMAGINE when additional analysis or data display will be necessary. Spatial data in the database, in general, will be pre-processed to allow compatibility in spatial referencing among different layers. Specialized viewing profiles such as application files for ArcView or viewing files for IMAGINE will be prepared for particular analysis or display purposes such as the viewing of LULC classification with a specific color scheme or an overlay of road buffers and the LULC map.

To facilitate data presentation, the project will also prepare hardcopies of image displays and geographic maps for selected data layers such as LULC maps and Landsat ETM+ imagery. Figure 3 is an illustration of the graphical display of the ETM+ imagery for a selected area along the I-405 corridor. These hardcopies will be first prepared in a standard softcopy image format (e.g., GeoTiff) to allow portability and to facilitate presentations (e.g., for inclusion in MS PowerPoint Presentations). Fundamental cartographic principles will be applied in the preparation of these hardcopies of images or maps that include proper use of map projections, map symbols, scales and annotations for effective data communications. Digital Elevation Model (DEM), when draped with remotely sensed imagery or feature data layers such as LULC maps, will be utilized to generate 3-D perspective views for more realistic visualizations of the area. There 3-D perspective views can be in the form of a hardcopy or softcopy. In the later case, ERDAS IMAGINE will be utilized to provide on-line 3-D viewing capabilities for the study area. That is, a user can choose or change a particular viewing angle, viewing direction or other viewing parameters while the 3-D scene is explored.

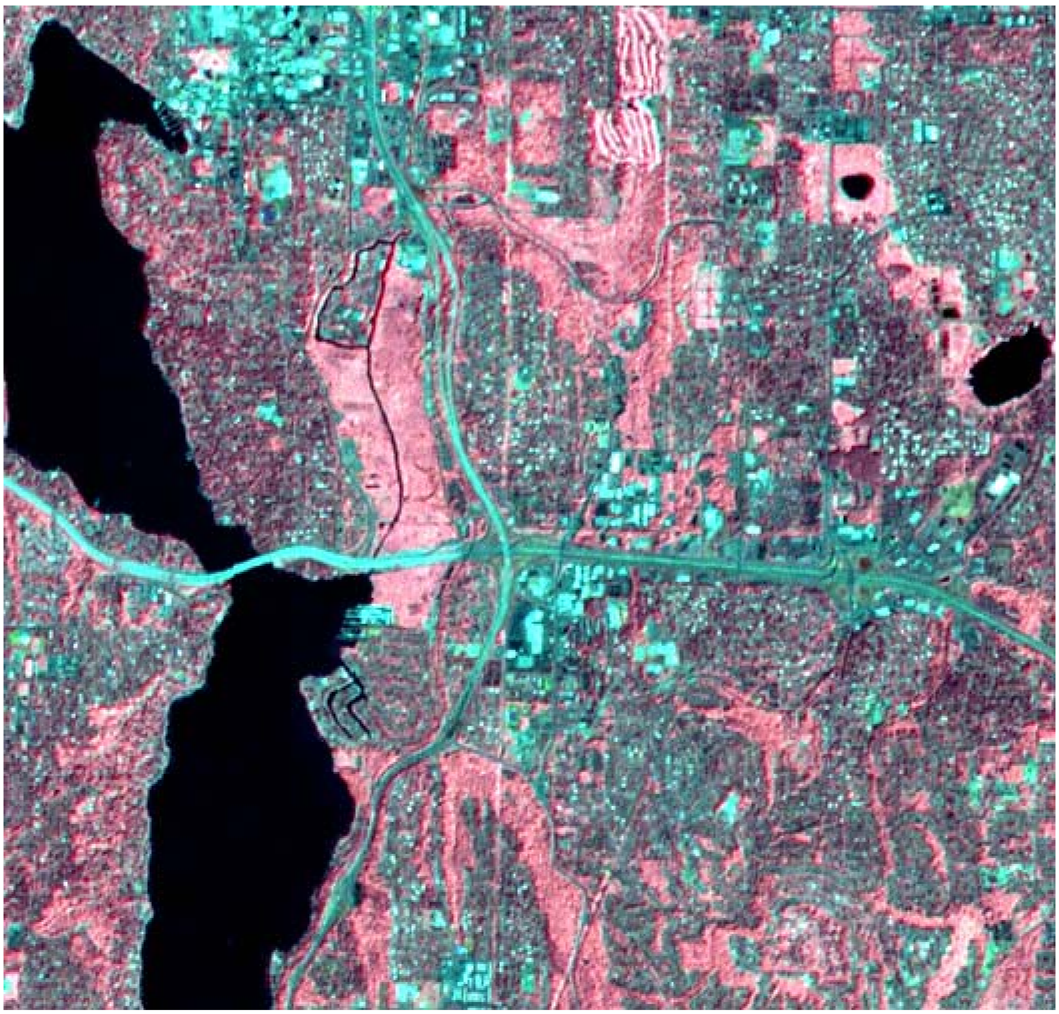

Figure 3. Landsat ETM+ Multispectral Imagery shows the intersection of I-405 and I-90. Graphical Display Generated with Landsat ETM+ Band 4 (Red), Band 3 (Green) and Band 2 (Blue) Sharpened to 15 Meter Resolution Using the ETM+ Panchromatic Band.

To qualify and quantify correlation between the proposed transportation infrastructure and the social and natural environments, a set of statistical tables will be prepared to summarize the social and natural characteristics along the corridor. Typically, LULC data will be characterized by land use types along the selected road segments and land use areas within a given distance from the selected road segments. Population and household counts will be established for a given buffer zone along the corridor. Water bodies as well as length of river streams will be measured in the neighborhood of the proposed infrastructure. Statistics obtained from data layers that represent feature characteristics in different time periods will be compared and differences will be identified to indicate changes and trends of changes.

References

Kiracofe, Bruce E., and Jerry Everett, Integrating Imagery Into Transportation Planning: Experience and Opportunities, originally copyrighted by the Transportation Research Board and reproduced in "GIS Technologies for the Transportation Industry", Urban and Regional Information Systems Association, Park Ridge, Illinois, 1999.

An Overview of the I-405 EIS Program, http://www.wsdot.wa.gov/I-405/overview/overview.htm (Accessed 10 Jan 2001).

Schwartz, Marcy, CH2M Hill, National Cooperative Highway Research Program – Active Project, Project 25-22, FY 1999 Technologies to Improve Consideration of Environmental Concerns in Transportation Decisions, 20 Sep. 2000, (http://www4.nationalacademies.org/trb/crp.nsf/All+Projects/NCHRP+25-22), Accessed 31 Oct. 2000.

Saulsbury, Bo, "Land Use/Land Cover Categories Used in the I-405 Draft EIS and Comparable USGS Classifications," Oak Ridge National Laboratory, May 2002.

ERDAS, *ERDAS Field Guide*, Fifth Edition, 2002.

ERDAS, *ERDAS IMAGINE Tour Guides*, 2002.

USEPA, *Indicators of the Environmental Impacts of Transportation*, EPA 230-R-96-009, October 1996.

Xiong, D., *Database Development of Land Use Characteristics along Major U.S. Highways*, ORNL/TM-2000/160, Oak Ridge National Laboratory, June 2000.